DOCUMENT RESUME

ED 308 461 CS 009 590

AUTHOR Feathers, Karen; And Others
TITLE Analyzing Retellings To Assess Comprehension.
PUB DATE Dec 88
NOTE 39p.; Paper presented at the Annual Meeting of the National Reading Conference (38th, Tucson, AZ, November 29-December 3, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Achievement Tests; College Freshmen; Grade 2; Grade 3; Higher Education; Holistic Evaluation; Primary Education; Reading Achievement; *Reading Comprehension; Reading Instruction; Reading Research; Text Structure
IDENTIFIERS *Story Retellings

ABSTRACT

Three studies examined the impact of different instructional reading materials and subsequent analysis of comprehension by retellings. The first study investigated the influence of the presentation of the material. Subjects, 83 freshmen enrolled in a biology course, were shown one of four combinations of course material involving a slide presentation and an accompanying taped presentation. Results indicated that traditional measures revealed no differences among the groups but that on the retellings the students were comprehending differently based on the context in which the information was presented. The second study examined the effect of text structure on retellings. Subjects, 56 third grade students, were provided instruction and later tested for ability to read and comprehend a passage containing three different text patterns--compare/contrast, enumeration, and sequence. Results indicated traditional measures did not reveal major trends in the reading process that retellings brought out. The third study investigated the effect of different instructional materials on children's reading behavior and comprehension. Subjects, 30 second grade students, read orally and then recalled four different types of stories immediately after reading. Results indicated a holistic scoring format allowed a consideration of aspects of comprehension not revealed by traditional information-based measures. The results of all three studies indicated that group mean scores tended to camouflage differences and that total scores did not reveal important information about students' comprehension. (Three tables of data and four figures are included; 39 references are attached.) (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED308461

Analyzing Retellings to Assess Comprehension

Karen Feathers, Virginia Reasor, Darrell Beauchamp

Symposium/paper presented at National Reading Conference

Tucson, Arizona

December, 1988

CS005590

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Darrell Beauchamp

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Analyzing Retellings to Assess Comprehension

When studying learning, whether from texts, various forms of media, or teacher instruction, the basic question is: What and how much did the student learn? This learning is typically assessed via a standardized or teacher-made comprehension test consisting of multiple choice, fill in the blank, and/or short answer questions.

When selecting comprehension measures, teachers and researchers alike face a dilemma. They can use multiple choice tests which can be given simultaneously to many subjects (thus saving time), can be scored quickly, and provide objective scoring (since only one answer is considered correct). However, what can be learned from this type of measure is limited by structured questions that require only "recognition" of important information as opposed to "recall" (Johnston, 1983; Farr & Carey, 1986). In essence, such tests measure not what students learned but whether or not they learned the specific information covered by the questions.

While there is certainly a need for assessing how much information students can recall, single answer questions are often used indiscriminately as the comprehension measure regardless of the type of text or the purpose for reading typically associated with that text or the context of situation. All too often, only one type of measure is used.

This can narrow the range of possible conclusions and produce misleading results.

An alternative comprehension measure, retelling, is accomplished by having students recall, either orally or in writing, everything they can remember from what they have read, seen, or heard. Oral retellings are time consuming because they are usually done individually, but even written retellings take time. Unlike objective tests, they allow the subjects to tell what they know about the topic, but they are more difficult to analyze and evaluate. Analysis is also time consuming and often involves subjective decisions about such things as matches between subject information and text information and the relative importance of information recalled. On the other hand, retellings allow for a variety of analyses that go beyond identification of text information and include organization, text structure, coherence, attitudes or opinion, relationships to other experiences, etc.

There are a variety of ways to analyze retellings. Kintsch (1974, 1977, 1980) and Frederiksen (1975) used propositional analysis and matched retelling propositions to text propositions. Van Dijk (1972, 1977, 1978, 1980) generated a text macrostructure. Meyer developed a text hierarchy and compared retellings to the text structure (1975, 1982). Retellings collected during miscue analysis (Goodman & Burke (1972) are compared with a "retelling format" that

summarizes text information, and Irwin and Mitchell have developed a holistic scoring system (1983).

Different conditions under which students received information, different texts, contexts, grade levels, ability levels, purposes for reading, and different research questions should lead to alternative approaches for analyzing comprehension and thus for analyzing retellings. Unfortunately, these factors are often not considered when selecting a method of analysis either in the classroom or in a research setting.

This paper will address the issue of the analysis of retellings by describing the methods of analysis used in three different studies and the conclusions that could be drawn from those analyses. In all cases both measures and analyses were tailored to fit the research, the subjects, the texts, and the contexts. The focus in this paper is not on the particular studies themselves but on the assessment measures and the subsequent analysis. Results of the studies will be used to demonstrate the appropriateness and effectiveness of the various measures.

The first two studies were similar in that both used an expository text as material and a text macrostructure as a basis for evaluating students' retellings. However, the procedures were not totally identical because the focal research question was different in each case. In the first study, the influence of the context of the message was

investigated, while in the second, the effect of teaching children specific text structures was assessed. The third study used narrative text instead of expository and a holistic scoring system to evaluate the retellings. The focus of this study was not context or teaching, but the impact of different instructional reading materials.

Analyzing Retellings to Determine the Influence of Context

This study investigated the variations in learning that might occur due to differences in the presentation of information. Today's children are exposed to a proliferation of media, televisions, movies, videos, computers, magazines, and billboards, which provide large amounts of visual stimuli (Heinich et al., 1985; Platt, 1975). In additon, slide-sets, with a recorded audio program on cassette tape, are widely used in school classrooms (Heinich et al., 1985; Kenny & Schmitt, 1981; Langford, 1973). These factors raise questions about the usefulness of the various forms of visual learning and their relationship to learning from print.

Since visuals appear to be so prominent in school and society in general, a linear slide + tape presentation approximately 7 minutes in length was used as the "text" for this study. Four variations of the same slide + tape presentation were developed consisting of the following: (Group PT) picture slides plus narration of text, (Group TT) print slides of text plus narration of text, (Group TM) print

slides of text plus music background, and (Group PM) picture slides plus music background.

Subjects were 83 students in four sections of a freshman level biological sciences course. At the beginning of the fall semester, a multiple choice test and an essay test were administered to determine prior knowledge and attitude toward the topic. Five weeks later, a slide + tape presentation was shown to each section. Immediately following the presentation, students completed a multiple choice test and a Likkert scale, answered an essay question, and recalled everything they could remember in writing. To minimize the effect of research procedures, all of the above were done as a normal part of the course of instruction.

Data Analysis

So that there would be a known variable with which to compare each student retelling, the script/text of the slide + tape presentation was parsed into individual concept/idea units and a seven level text macrostructure (Rector, 1987; van Dijk, 1977; Vaughan, 1984) was developed consisting of: Gist, Superordinate Levels 1 and 2, Subordinate Levels 1 and 2, Nonessential Details Levels 1 and 2.

Retellings were also parsed into idea units and compared with the narrated text macrostructure. Retelling units stating explicit information from the text were given a weighted rating depending on the level in the macrostructure (Vaughan, 1984; Rector, 1987). For example, a gist statement was given

30 points, superordinate ideas 25 and 15 points for levels one and two respectively, and subordinate ideas 10 and 5 points for levels one and two. Retelling units generated by the reader on the basis of text information (implicit) were given the weighted value of the related text explicit information. Text general statements, statements about the presentation, and non-essential details were given values of 3, 2, and 1 point, respectively. In addition, each retelling was evaluated by considering coherence, clarity, support for major ideas, and inclusion of a gist statement.

Because this type of analysis is subjective, all procedures were done independently by three raters and the results compared. Agreement by two of the three raters was accepted.

Discussion of Results

Although these findings are not obvious in the initial data anlysis, there were differences manifest depending on the form of presentation of information, not only in the information learned, but also in what the students focused on and chose to talk about. Crosstabulations and one way analyses of variance were used to consider differences among the four groups, and only the mean score of Group PM (27.21) differed significantly from Group PT (93.33), Group TT (93.68), and Group TM (86.76).

However, there are many ways to achieve a total score. For instance, a few high level subordinate and/or superordinate statements could produce the same total score as

many details or details plus a gist statement. Additionally, mean scores do not reveal important variations within groups. Therefore, neither total scores nor a comparison of mean scores reveal features that distinguish these groups.

Even when scores were not significantly different, crosstabulations disclosed characteristics of the four groups that the traditional multiple choice test and consideration of the total retelling scores did not. For instance, this analysis indicates that Group TT retellings consistently contained a greater number of words than did the other three groups, suggesting that a heavy dose of cognitive stimuli (narrated text plus text on the screen) prompted students to write more. These retellings also had more comments clustered at the top of the macrostructure, but contained a larger number (27%) of false statements. Students in Group TT were able to repeat more of the information from the presentation, but, in many cases, that information was incorrect. Retelling #23 (Figure 1) is an example of retellings from this group.

Figure 1 Here

The retellings for Group TM call into question the findings of the multiple choice test which indicate that this

group had the highest mean score for that test. An analysis of the Group TM retellings identified two clusters of retellings, one with comments that matched high level text items and the other with comments that matched low level items. The cluster in the high macrostructure levels indicates that there were some students able to organize the minimal "reader slide only" information (Figure 1, Retelling #50). At the same time, however, the cluster of retellings in the lower macrostructure levels suggests that there were students in the same group who were not able to comprehend and organize this information (Figure 1, Retelling #49). These two clusters were also obvious in the scores for coherence, clarity, and organization, with 42% of the retellings rated low and 38% rated high. Additionally, a large percentage (48%) of these retellings contained no comments at superordinate level 1.

Group PT retellings contained more implicit details than any other retellings. This would suggest that the pictures in conjunction wih the text prompted students to relate the information being presented to their own experience. Retellings in this group also were longer with 40% being in the top 20% of total number of words. A high percentage of retellings (60%) included comments at high levels of the macrostructure (superordinate level 1, 60%; superordinate level 2, 47%), and more retellings included multiple comments at this level. Finally, a large number of retellings for this

group were in the upper 50% of total scores (67%), with 40% in the upper 20% of total scores (Figure 2, Retelling #4).

Figure 2 Here

Although it was obvious from the traditional multiple choice test and the retelling total scores that the students in Group PM did not do as well as students in the other three groups, it was the retellings that offered insight into what was happening within that group. Group PM retellings contained few cognitive facts, and, in all but a few cases, were centered around a listing of the pictures in the presentation (Figure 2, Retelling #68). In contrast, retellings of Groups PT, TT, and TM were centered on the cognitive information. Retellings for Group PM also contained fewer words and received low scores for clarity, organization, and coherence.

Clearly, very little information was contained in these retellings. Since no cognitive information was communicated through language, there was little that students could write about. Indeed, it is interesting to note that students dealt with this dilemma of being required to retell by listing what had been viewed while only two retellings contained comments

about the music. Apparently, the pictures were considered more valuable information than the music.

Group TM had the highest mean scores on the multiple choice test, but analysis of the retellings indicates that Group TM as a whole did not produce the best retellings. In actuality, the retellings fell into two clusters, with some students doing very well while others did poorly. The retelling total scores indicate that Group TT was highest in terms of recall of information. However, since a large percentage of this information was false, Group PT appears to have the most correct information. Additionally, unlike Groups TT and TM, retellings of Group PT contain many implicit details. This finding is important since the ability to infer from one's experience and the text information is an important aspect of comprehension.

Conclusions

Thus, although traditional measures did not indicate any differences among these groups, it is clear from looking at the retellings that the students were comprehending differently depending on the context in which the information was presented. Because mean scores mask differences, the mean retelling scores were not as informative as the analysis of the retelling patterns within each group. Only this deeper analysis revealed the subtle differences that distinguished these groups.

Analyzing Retellings to Determine the Influence of Text Structure

Research in recent years suggests that awareness of text structure is an important factor of comprehension of informational material, and various means of helping students become aware of text structure have been investigated (Anderson, 1978; Armbruster, 1980; Meyer, 1975). Most often, some form of direct instruction of particular structures such as cause/effect, comparison/contrast, time sequence, etc., are used with a group of students and then the subjects are given some form of comprehension test. This study is similar in that 30 third grade subjects were provided instruction and later tested for ability to read and comprehend a passage containing three text patterns - comparison/contrast, enumeraton, and sequence. However, direct instruction was not used. Fifty-six third grade students were assigned to experimental and control groups on the basis of reading level and gender. The experimental group was engaged in writing passages based on these patterns and in mapping texts with these patterns, and the control group did worksheets on sequencing and comparison/contrast. Two weeks after the instruction, subjects read a complete subsection of a chapter in a third grade science book. Immediate written retellings were collected.

To test the hypothesis that the text organizational pattern served as a cue to process the author's intended

message, a measure was needed which would reflect the use of text structures by each subject and how this use affected comprehension, including reader-based information as well as text-based information. Since traditional comprehension measures could not meet these criteria, a free written retelling was used.

Data Analysis

Two experts independently parsed the text into concept units and generated a macrostructure (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1977). The two were then compared and a final macrostructure consisting of five levels was agreed on: Gist, Superordinate Level 1, Subordinate Levels 1 and 2, Nonessential Details. Concept units were given a weighted value determined by the level of the unit in the text macrostructure (Beauchamp, 1988; Rector, 1987; Vaughan, 1984). Each retelling was also parsed, and, to allow a comparison of text structure, a macrostructure was developed for each retelling. All procedures were done by two independent raters. Disagreements were resolved by discussion.

Retellings were then considered from three perspectives: comparisons between the text macrostructure and readers' macrostructures, identification of text patterns, and overall quality of retelling. First, concept units were categorized according to the source of information (text explicit, text implicit, reader-based). Then, to allow a comparison of quality of information as well as quantity, each concept was

given a weighted score according to the level of the concept in the macrostructure. Concepts appearing high in the macrostructure received a greater value than those appearing low in the macrostructure. These weighted values were summed to produce a total retelling score.

Finally, since the primary focus of this study was to determine the effects of teaching text structure, text patterns in retellings were compared with the original text patterns. The pattern was identified (comparison/contrast, enumeration, sequence), and then rated on a scale of 0 to 4 according to how much information from the original text pattern the retelling contained.

Discussion of Results

A one-way analysis of covariance with pretest reading comprehension scores as the covariate and Reading Total as the dependent variable indicated a significant difference between the experimental group and the control group ($F=4.25$, $p.<.05$). Mean score for the experimental group was 92.82, while that for the control group was 68.46.

Ratings of text structure revealed differences between groups not identified by the total retelling score. For the purpose of analysis, text structure scores were collapsed into two categories. Scores ranging from 0-2 indicated little or no use of text structure. Scores ranging from 3-4 were considered good to extensive use of text structure. Twelve students (43.2%) from the experimental group made extensive

use of structure as opposed to seven (25%) from the control group. Twice as many subjects in the experimental high group as compared to the control high subjects made extensive use of text structure in their retellings. There was no difference in the low groups.

In addition, the use of text structure had an effect on overall quality of retelling. Students who made little or no use of text structure received the lowest scores of 0 or 1 for coherence, organization, and clarity, while those who utilized text structure received the highest scores. The same trend occurred with Total Reading scores. Eighty-one percent of the subjects not utilizing text structure placed in the lowest 40% of the scores, while 52.6% of those students using text structure extensively placed in the top 40% of the scores. Additionally, there were no gist statements included for any subject not utilizing text structure, but of those students who did employ structure, 41.1% included a gist statement.

For instance, retellings #54 and #41 (Figure 3), retellings that contain text structure patterns, include gist statements (There are many kinds of stems) and the comparisons of stems found in the original text. Retelling #54, written by a high ability student, also contains an enumeration pattern and is longer than retelling #41. However, this retelling, written by a low ability student, does suggest that this type of instruction can be effective, although retellings of low students as a whole did not contain the text

structure patterns. It might be that the length of instruction needs to be increased for these students so that they have more experiences mapping and writing.

Figure 3 Here

By comparison, retelling #3 (Figure 3), from a high control group student, and #20 (Figure 3), from a low control group student, do not contain gist statements or structure patterns. Retelling #3 is a listing of topics discussed in the original text, and retelling #20 is a listing of details. Both appear to have no organization; the topics/details chosen are not arranged in any particular order. Retelling #3 refers to the sequence patterns of food distribution and plant growth, and the comparisons of different kinds of stems, but no details are provided on any of these topics. Retelling #20 contains no evidence of text patterns.

Conclusions

The data from this study underscore the complexity and power of the reading process. The traditional measures are important, but only allow comprehension to be judged from a perspective of total score and did not reveal other major trends. In this case, free retelling analysis tailored to fit the situation allowed an exploration of comprehension from

many perspectives.

Analyzing Retellings to Determine the Influence of Text Type

The last study was concerned with the effect of different instructional materials on children's reading behavior and comprehension. Researchers (Altwerger & Goodman, 1981; Donaldson, 1978; Harste, Burke, & Woodward, 1981), posit that school instruction in reading must build on the child's language data base and specifically that reading material must reflect the child's language. Others suggest that differences exist between the language used in basal readers and the language of the child (Sampson, 1972), and that differences in texts affect reading (Rhodes, 1979; Thomas, 1980). With the preponderance of materials now available for use in reading instruction, both basal readers and children's literature, it behooves us to determine if and how different materials affect comprehension.

To investigate this issue, 30 second grade subjects each read orally and then recalled four different types of stories - a story from a skills basal reader, one from a phonics basal reader, a trade story (from Bill Martin's Sounds of Language) and a student dictated story. The readings of the various types of texts were randomized, and all retellings occurred immediately after the oral reading. Both the oral readings and the retellings were tape recorded for additional analysis.

Data Analysis

The method of analyzing the retellings in this study is different from the usual method of assessment based on some

form of a matching of "text" information with "reader" information, for example propositional matching (Kintsch, 1974) or Meyer's (1975, 1982) hierarchy of idea units. It grew out of other methods of assessment, which attempt to consider the retelling as a text, for example Irwin & Mitchell (1983). However, it is different from these holistic assessments in that it allows consideration of each category independently of others.

Current holistic assessments are based on a theoretical view of what a poor vs a good retelling is like and a series of levels is established, each of which contains all the elements of the lower levels plus new elements. For example, Irwin & Mitchell's level 3 contains: major ideas, appropriate supporting details and relevant supplementations, adequate coherence, completeness, comprehensibility. Level 4 includes all of these plus a thesis statement. The problem with this type of holistic scoring is that retellings often do not fit the descriptions for the various levels. For instance, a student retelling "It was about an old woman who tried and tried to sell her jelly, and finally she succeeded," contains the gist statement that occurs only in level 4 or 5 of the holistic scoring format, but does not contain the major points, supporting details, etc., that are indicated for a score of 4 or 5. The scoring procedure used in this study eliminates this problem by rating each descriptive category separately.

Additionally, the categories were not predetermined but developed from the data, using retellings from incomplete data

sets. First three raters independently read 15 retellings, separated them into three groups (poor, acceptable, good), and ranked them from poorest to best within each group. Then the raters compared their groupings and rankings, and, based on their discussion of the reasons for their placements, developed a list of descriptive categories (Figure 4).

Figure 4 Here

Since information remembered from texts is an important factor of this type of evaluation, two categories (1 & 2) assess the amount of text information included in the retelling. Five categories (3 - 7) have to do with assessing the retelling as a text in itself: (3)the inclusion of a theme or topic, (4)to what extent the presentation of events is logical, (5)to what extent the teller relates the events in complete statements as opposed to fragments of thought, (6)to what extent the retelling makes sense, (7)to what extent the retelling includes conventions of story telling. Finally, since reading narrative means more than simply recalling information, four categories evaluate the extent to which the retelling provides evidence that the reader noted particular literary aspects of the original text (8 -11), including dialogue, stylistic patterns, text wording, and text structure.

The eleven categories generated in the initial analysis were used to evaluate the remaining complete sets of retellings (9 sets of 3 = 27 retellings). Three raters read each retelling and rated each category on a scale of 0 to 3 with zero indicating no evidence from that category in the retelling and three indicating a high amount of evidence from that category in the retelling. The score used in final analysis was the score assigned by at least two raters. Ratings for two categories considered of primary importance were doubled when summing for total score. Ratings for all eleven categories were totaled for a retelling score.

Discussion of Results

Both analysis of variance and crosstabulations indicated differences in comprehension depending on the type of story that was read (Feathers & White, 1987), findings similar to Rhodes, 1979) and Thomas (1980). However, of more importance were the differences produced by using this type of evaluation as compared with a more "information based" form.

To illustrate the difference between an information based score and a holistic score, retellings were scored using the retelling format for the Reader Miscue Inventory (Goodman & Burke, 1972) as the information based score. A RMI story format identifying the major details from the story, plot, gist, theme, characters, and character description was developed for each story and points were assigned for each component. Retellings for each story were compared with the RMI format, and points for components were totaled. To allow comparison, both old and new scores were converted into percents by dividing the raw score for each student by the total possible.

A comparison of the old and new mean retelling scores (Table 1) indicates fewer differences for the average group, than for the low and high groups. These mean scores, however, do not reflect what happened with individual scores and/or what is not learned about students when only the information based score is considered. The comparison of individual scores indicates differences among some scores (Table 2) that are not represented by the mean scores.

Table 1 Here

Table 2 Here

Further, consideration of the categories of the retelling evaluation format reveal additional information about students' comprehension not provided by the total score (Table 3).

Table 3 Here

Four retellings will be used to demonstrate some of these differences and thus illustrate what a researcher/teacher might or might not learn about students depending upon what scoring system is being used to evaluate retellings.

Retelling #28 (Figure 5), for example, scored lower on the new format than on the old (Table 2). The higher old score is primarily due to the inclusion of the major characters (Kid Colorado, Mom, and Goldenrod), and the inclusion of three major events, mom didn't bounce him on her knee, Kid Colorado made more pasture, and Kid Colorado's girlfriend comes along on Goldenrod.

Figure 5 Here

This retelling contains some important details but little else. For example, there is no topic statement, logical order, dialogue, or repetition of stylistic pattern or text structure. While neither the old or the new score is very high, much more is learned from the new holistic scoring than the old scoring in terms of what this student was or was not aware of.

Retelling #4 (Figure 5), from a student in the high group, includes a lot of information. This retelling essentially reiterates all the major events in the story. The old scoring, however, give no credit for the inclusion of dialogue, recognition of text structure, and the logical ordering of events evident in this retelling, all of which are indications of comprehension (Figure 4).

Retelling #16 (Figure 5) is a clear example of a good retelling that doesn't score very high on the old format because not enough details are included (Figure 3). Each major episode in the story is summarized in this retelling, but since no details are included, the old score is low and is not a true reflection of the student's comprehension. With the new format, the student's understanding of the basic structure and theme can be credited as well as those qualities of logical ordering, complete thoughts, and coherence (Figure 4) that make this a good retelling.

Finally, two retellings from the same student in the low group indicate that information based scoring does not reveal

what students know. Retelling #9b (Figure 6) is obviously longer than #9a (Figure 6), but the amount of information included is not the only difference between these two retellings. Retelling 9a identifies the major character and then tells some things about him, but it is unorganized and jumbled to the point that anyone not familiar with the story would not be able to recognize it from this retelling. There are, however, some saving graces in this retelling. The topic is identified (Herbert), some major events are referred to, and an aspect of story telling is evident in the beginning, "It was about Hebet", and the ending, "That's all folks". These all indicate that this child is aware of certain aspects of stories and retellings (Figure 4).

Figure 6 Here

With a different story on the same readability level, the retelling looks quite different. The added length of 9b is indicative of the greater amount of information included as reflected in the increase in the old score from 11% (9a) to 36% (9b) (Figure 3). Nevertheless, this increase is not indicative of other, more important, differences between these retellings. Retelling 9b, unlike 9a, is organized, logical, and coherent. The basic text structure is followed, and dialogue and stylistic patterns present in the original story are found in this retelling (figure 4).

A teacher or researcher looking only at the initial retelling or considering only the information based score would probably conclude that this student needed a great deal of help with some basic comprehension skills. Retelling 9b, on the other hand, indicates that this student has the ability to understand and organize story sequence and is also aware of other aspects of stories such as structure, dialogue, literary patterns, etc.

Conclusions

This holistic retelling format allowed a consideration of aspects of comprehension not revealed by the information based measure. Comprehension, especially the comprehension of literature, is more than a gathering and repetition of information; therefore, assessment of comprehension needs to consider these other aspects.

The format described here allows these other elements to begin to be considered. That is not to say that this scoring system is ideal. In fact, there are some major omissions that need to be addressed. For instance, one problem is establishing a way of considering the reader's reactions to the text, i.e., like or dislike, comments about characters, events, etc. that indicate a true interaction with the text. A second problem is how to include credit for connections to their own lives that readers might make, such as, "I had a dog like that once," or, "Mrs. Purdy reminds me of my grandmother," or, "I don't understand how he could do that in the dark closet." A third problem area is how to give credit for reader response that indicates an involvement with characters and events. For example, one retelling included "WAHOOOO!" said very loudly three times. This child was clearly responding to the text in a very personal way that in not considered in current evaluation systems including the one proposed here. On the other hand, the format suggested here can be expanded by additional categories, and thus may provide a basis for consideration of total comprehension.

Summary

Three alternative means of evaluating retellings were described in this paper. While each was different, some similarities do exist. For example, all three include a consideration of the amount of information recalled; each considers importance of information as well as quantity; and all evaluate the retelling as a text with some measure of

organization, coherence, and/or clarity.

Because all of these measures are considered subjective, at least two raters were used for all judgements. Although this would be recommended for any research using these forms of evaluation, it would not be necessary when used for instructional purposes since rater agreement never fell below 96%.

In spite of the fact that these means of evaluation provide a "total retelling score", much more was learned in all these analyses about students' comprehension when separate categories were examined and individual and/or group patterns were explored. Group mean scores tended to camouflage differences, and total scores did not reveal important information about students' comprehension.

While each of these analyses provides a different means of evaluating retellings, they are all considered "working models" of assessment because there are many facets of comprehension either not tapped or unsatisfactorily evaluated by these measures. For instance, in the first two studies, attempts were made to consider inferences, either text-based or reader-based, but these statements were evaluated in comparison with the text macrostructure. A macrostructure was also developed for each retelling, but the problem of how to evaluate that macrostructue has not yet been resolved.

In addition, factors such as: connections the reader makes between personal experiences and text information, reader reactions to information, characters, actions, etc., comments about the text itself, evaluative comments about

content or written quality of the text, etc., are not considered with these assessments. These methods, however, do provide forms of analysis that allow multiple views of data. Further, since these formats are open and flexible, there is room for expansion and/or alteration that may eventually allow a more complete consideration of the total process of comprehension.

REFERENCES

Altwerger, B. & Goodman, K.S. (1981). Studying text difficulty through miscue analysis (Occasional paper No. 3). Tucson: University of Arizona, Program in Language and Literacy.

Anderson, T.H. (1978). Study skills and learning strategies (Tech. Rep. No. 104). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Armbruster, B. E. (1980). The effect of mapping on the free recall of expository text (Tech. Rep. No. 160). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Beauchamp, D. (1988). The effects of slide + tape presentation type on affect of response and cognitive achievement. Unpublished Doctoral Dissertation. East Texas State University, Commerce, Texas.

Donaldson, M. (1978). Children's Minds. New York: W.W. Norton & Company.

Feathers, K.M. & White, J.H., (1987). Reader reactions to text. Paper presented at the annual conference of the American Educational Research Association, Washington, D.C.

Frederiksen, C.H. (1975) Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology. 7. 371-458.

Goodman, Y. & Burke, C. (1972). Reading miscue inventory. New York: MacMillan.

Harste, J.C., Burke, C.L., & Woodward, V.A. (1981). Children their language and world: Initial encounters with print (Final Report). Washington, D.C.: National Institute of Education.

Heinich, R., Molenda, M., & Russell, J.D. (1985). Intructional media and the new technologies of instruction (2nd ed.). New York: John Wiley & Sons.

Irwin, P.A. & Mitchell, J.N. (1983). A procedure for assessing the richness of retellings. Journal of Reading, 3, 391-396.

Kenny, M.F., & Schmitt, R.F. (1981). Images, images, images (2nd ed.). Rochester, N.Y.: Eastman Kodak.

Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, New Jersey: Erlbaum.

Kintsch, W. (1977). On comprehending stories. In M. Just, & P. Carpenter (Eds.) Cognitive Processes in Comprehending. (33-63). Hillsdale, New Jersey: Erlbaum.

Kintsch, W. (1980). Learning from texts, levels of comprehension or: Why anyone would read a story anyway. Poetics. 9 87-98.

Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. Psychological Review. 85. 363-394.

Langford, M.J. (1973). Visual aids and photography in education; A visual aids manual for teachers and learners. New York: Hastings House.

Meyer, B.J. (1975). The organization of prose and its effect on memory. Amsterdam: North-Holland.

Meyer, B.J. (1982). Prose analysis: Purposes, procedures, and problems, part II (Res. Rep. No. 11). Tempe: Arizona State University, Department of Educational Psychology.

Platt, J.M. (1975) Visual Literacy. Washington, D.C.: National Education Association.

Rector, B.S. (1987). A comparison of reading comprehension profiles among students with analytic, global, or integrative cerebral dominance. Unpublished doctoral dissertation, East Texas State University, Commerce, Texas.

Rhodes, L.K. (1979). Comprehension and predictability: An analysis of beginning reading materials. In J. Harste, & R. Carey (Eds.). New perspectives on comprehension. (100-131). Bloomington, Indiana: Monograph in Language and Reading Studies.

Sampson, M. A (1972). A comparison of complexity of children's dictation and instructional reading materials. In J.A. Niles & L.A. Harris (Eds.) 31st Yearbook of National Reading Conference. (177-179).

Thomas, S. (1980).The effect of reader-authored materials on the performance of beginning readers. Unpublished Doctoral Dissertation, Michigan State University. 41. #11, p.4669A.

Turner, A. & Greene, E. (1977). The construction and use of a propositional text base (Tech. rep. No. 63). Boulder: University of Colorado, Institute for the Study of Intellectual Behavior.

van Dijk, T.A. (1972). Some aspects of text grammars. The Hague: Mouton.

van Dijk, T.A. (1977). Semantic macrostructures and knowledge frames in discourse comprehension. In M.A. Just, & P. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum.

van Dij,, T.A. (1978) Recalling and summarizing complex discourse. In W. Burghardt & K. Holker (Eds.). Text Processing. (pp. 49-118). Berlin: de Gruyter.

van Dijk, T.A. (1980). Macrostructures. Hillsdale, New Jersey: Erlbaum.

van Dijk, T.A., & Kintsch, W. (1977). Cognitive psychology and discourse: Recalling and summarizing discourse. In W. Dressler (Ed.), Trends in textlinguistics. (pp. 61-80). Berlin: de Gruyter.

Vaughan, N.E. (1984). A comparative evaluation of reader's comprehension as measured by free hierarchially structured questions (Doctoral dissertation, East Texas State University. 45/#12A, p.3597A.

Table 1

Old and New Mean Retelling Scores

Group	Old Scoring			New Scoring		
	Phonics	Skills	Trade	Phonics	Skills	Trade
Low	23	9	30	16	13	32
Average	13	37	21	11	38	19
High	39	32	28	42	34	64

*Scores are in percents

Table 2

Comparison of Old and New Retelling Scores

Subject	Old Score	New Score
28	21	13
16	34	49
4	31	74
9a	11	23
9b	36	59

*Scores are in percents

Table 3

New Retelling Scores by Categories

	Subject				
Retelling Category	28	16	4	9a	9b
Major events	1	2	3	1	2
Explication of events	1	0	2	1	2
topic	0	4	4	4	4
Logical order	0	2	3	0	3
Complete thoughts	0	3	2	1	2
Coherence	0	6	6	0	4
Story telling	2	1	0	1	0
Dialogue	0	2	3	0	1
Stylistic pattern	0	0	2	0	1
Text wording	1	0	1	0	2
Text structure	0	2	3	0	2
Total	5	21	29	9	23
Total Percent	13	54	74	23	59

Figure 1

Sample Media Retellings

Retelling #23 (Group TT)
The Sargasso Sea has the most large plants. There is only one plant that can live in large bodies of water without being attached to the surface, phytoplankton. It can reproduce rapidly, therefore, it's population increases. Living organisms live on the nutrients of phytoplankton. Estuaries als bring nutrients to water surfaces.
About 10% of the fish population is taken by man for food. The herbivores are tiny because of the smallness of marine plants. Larger ocean animals are carnivores and are confined to eating mostly phytoplankton. Large fish are scarce. There in not much food in the ocean, so man must depend on land to get most of their food.

Retelling #50 (TM High)
Ocean farming is not feasable since there is no way to fertilize the ocean. The ecosystem is so closely interdependent and cannot be changed to produce more than it presently does. Limiting factors include the facts that surface waters are poor in nutirents, light does not reach the bottom depths, large plants could hoard nutrients. Phytoplankton are the only plants that can maintain their population. Their smallness gives them a high effectiveness. Oceans make up 3/4 of earths surface but produce only 1/4 of its calories. 100 million tons of fish could be taken from the ocean without hurting stock. 60 million are already being taken.

Retelling #49 (TM Low)
The slide presentation talked about the chance of farming the ocean in the future. The ocean produces 140 million tons of food. Man now uses 60 million tons. If he uses 100 million he will ruin the stock of food. 3/4 of the earth's surface is covered with water. Rich fisheries are 10% of the ocean.

Figure 2

Sample Media Retellings

Retelling #4 (PT)
The film based on ocean or marine life focused in on where and how different plants and animals survive in ocean waters. Phytoplankton is the basic plant in the food chain which survives in the sunny part of the water and stores the nutrients in its leaves. Phosphorous is necessary to plant life and plants are necessary to animal life. While some animals feed on other animals there must still be plants for survival. Most plants grow in shallow water in the shore where they can attach themselves to the floor to keep from being carried away from the tide. In the case of phytoplankton, its nutrients are sometimes washed out to the deeper parts of the ocean for animals to use in their struggle for survival. The most productive place for animals is the estuaries where salt and pure water combine and nutrients are rich. There is even a possibility for plants and animals to srvive in cold anarctic (arctic) waters. The animals in the water are limited to and we probably won't be able to grow more than what the waters already produce. Therefore we aren't for sure what will happen to our future ocean supply of food and we will probably need to depend on the land for our food.

Retelling #68 (PM)
The first thing I saw was a man in a rubber raft, perhaps an oceanographer studying the plant and animal life in the ocean, there was a great deal of plankon floating in the water on one slide. Also people on the beach, a whale with two smaller whales, something that looked like a sand dune, or desert, man pouring water while in raft. 3 penguins on a ice cap, fishing boat with seabirds flying above, fish that looked like it was transparent, small odd shaped specimens, basket of small fish on a dock with three men near by.

Figure 3

Sample Science Retellings

Retelling #54
There are many kinds of stems Cattail stems are hollow pea vins twist and curve strawbery plants grow along the gound the stem of a tree is very large the stems cary menrals to all parts of the plant. some stems grow under ground like a potato buds are the flower before it blooms. it starts growing in the summer and stop in winter and grow more in winter sometimes you can plant the pulb of the plant and it will grow like a potato plant.

Retelling #41
There are many kinds of stems on trees. One of the stems have buds onion are like that. Some stems curve strawberry trees are like that. Some parts of a tree curry minarals fram one part to another. Some stems grow underground. pataes are like that.

Retelling #3
It was teaching you about how stems get water. And how they grow. It was teaching about woody stems and all kinds of stems. And that you can plant eyes of potatoes and the bulbs of unins. And some trunks are hollow.

Retelling #20
Woody stems can be found in the woods and parks. A solf stems is almost always green. potoe groys from its eyes. Some plant start all over ever year. Pee plans are hollow in the midle. onions groy by there bults. Some stems stay onder the grand all the time.

Figure 4

Categories of Retelling Evaluation

1. INCLUSION OF MAJOR EVENTS
2. EXPLICATION OF MAJOR EVENTS
3. TOPIC STATEMENT
4. LOGICAL ORDER
5. COMPLETE THOUGHTS
6. COHERENCE
7. STORY TELLING
8. DIALOGUE FROM TEXT
9. INCLUSION OF A STYLISTIC PATTERN FROM THE TEXT
10. INCLUSION OF WORDING OTHER THAN DIALOGUE
11. REPETITION OF TEXT STRUCTURE